Cytological Features of Peanut Genome

4

Guillermo Seijo, Sergio S. Samoluk, Alejandra M. Ortiz, María C. Silvestri, Laura Chalup, Germán Robledo and Graciela I. Lavia

Abstract

This chapter aims to update the chromosomal features evaluated by classical and molecular cytogenetic techniques. Karyotype variability detected within and among species was very useful to unravel the taxonomy of the genus and to establish relationships among species. This chapter includes analyses of chromosome morphology, heterochromatin, rDNA loci, as well as dispersed and clustered repetitive sequences. A critical review of the genome sizes of *Arachis* species is also provided. The usefulness of chromosome data is presented in three examples. The first one deals with the origin of the cultivated peanut. Molecular cytogenetics evidenced that the varieties of *A. hypogaea* may have had a single genetic origin, that *A. monticola* is a direct tetraploid ancestor of peanut, and that *A. duranensis* (A genome) and *A. ipaënsis* (B genome) are the diploid progenitors of the AABB tetraploids. The second one pointed to the analysis of the origin of the rhizomatous tetraploids and their relation to the unique diploid species (*A. burkartii*) of section Rhizomatosae. The cytogenetic data suggest that *A. burkartii* has to be discarded as a genome donor of the tetraploids, and that the latter may have had independent origins involving different species. The third one concerns the species of section Arachis, and how the chromosome data aided in the establishment of the genome groups (A, B, D, F, G, and K).

G. Seijo · S.S. Samoluk · A.M. Ortiz · M.C. Silvestri · L. Chalup · G. Robledo · G.I. Lavia (✉)
Instituto de Botánica del Nordeste (CONICET-UNNE), Sargento Cabral 2131, C.C. 209, 3400 Corrientes, Argentina
e-mail: graciela.lavia@yahoo.com.ar

G. Seijo · A.M. Ortiz · G. Robledo · G.I. Lavia
Facultad de Ciencias Exactas y Naturales y Agrimensura, Universidad Nacional del Nordeste, Av. Libertad 5460, 3400 Corrientes, Argentina

4.1 Introduction

The genus *Arachis* is composed of 81 species (Krapovickas and Gregory 1994; Valls and Simpson 2005; Valls et al. 2013; Santana and Valls 2015) distributed within a large region of South America, which extends from the eastern foothills of the Andes Mountains in Bolivia and

R.K. Varshney et al. (eds.), *The Peanut Genome*, Compendium of Plant Genomes, DOI 10.1007/978-3-319-63935-2_4

northern Argentina to the Atlantic coast of Brazil, and from the southern limit of the Amazonian rainforest toward the northern coast of La Plata River in Uruguay. Based on morphology, cross-compatibility, viability of the hybrids, geographic distribution, and cytogenetics, the *Arachis* species have been arranged in nine taxonomic sections: Trierectoides, Erectoides, Extranervosae, Triseminatae, Heteranthae, Caulorrhizae, Procumbentes, Rhizomatosae, and Arachis (Krapovickas and Gregory 1994; Fernández and Krapovickas 1994; Lavia 1996; Valls and Simpson 2005). Cross-compatibility, karyotypic, and meiotic analyses also allowed the identification and description of six different genomes within the section *Arachis*, namely, A, B, D, F, G, and K (Smartt et al. 1978; Stalker 1991; Robledo and Seijo 2008, 2010; Robledo et al. 2009; Silvestri et al. 2015). The genomic constitution of the remaining species of the genus, in the absence of comprehensive cytogenetic and molecular analyses, is less precise and has been traditionally assigned to different genomes on the basis of the subgeneric divisions, that is, Am (Heteranthae), C (Caulorrhizae), E (Trierectoides, Erectoides and Procumbentes), Ex (Extranervosae), T (Triseminatae), and R (Rhizomatosae) (Smartt and Stalker 1982).

Classical and modern molecular cytogenetics revealed a huge variability within and among species of different sections. These studies provided important information about the complexity of the peanut genome and were very useful to unravel the taxonomy of the genus and to establish relationships among the wild species and between them and the cultivated peanut. Here we present an update of the cytological information on *Arachis* species and some examples in which the use of chromosome markers was decisive to understand critical and long-lasting problems in the genus.

4.2 General Chromosome Features

The available information evidences that the genus *Arachis* is dibasic, with a great predominance of the species with $x = 10$, and only four species with $x = 9$ (Fernández and Krapovickas 1994; Lavia 1996, 1998; Peñaloza et al. 1996; Peñaloza and Valls 2005). Most of the species of the genus are diploids, except the tetraploids *A. hypogaea* and *A. monticola* of section Arachis, three tetraploid species of section Rhizomatosae, and one triploid accession of *A. pintoi* of section Caulorrhizae.

Chromosome identification in peanut began with the analyses done by Husted (1933, 1936), who distinguished one pair of chromosomes that borne an extended secondary constriction (B chromosome pair, SAT chromosomes after Fernández and Krapovickas 1994) and another one that was conspicuously smaller than any of the others of the complement (A chromosome pair, A9 after Robledo et al. 2009). The karyotypes of peanut and *A. monticola* are highly symmetric with a predominance of metacentric chromosomes. The most common karyotype formulae are 38 m + 2 sm and 36 m + 4 sm. Except the A9 pair, the other chromosomes of the karyotype are very similar in size (1.88 μm on average), whereby the chromosomes of the A and B genomes are indistinguishable by classical techniques (reviewed in Fernández and Krapovickas 1994; Lavia and Fernández 2004). Cytogenetic studies covering the six taxonomic varieties of peanut did not reveal major karyotype differences among the subspecies *hypogaea* and *fastigiata* (Lavia and Fernández 2004). Only one (very rarely two) pair of SAT chromosomes has been usually distinguished in all the varieties and in *A. monticola*. However, different types of SAT chromosomes, types 3, 5, and 6, according to the classification of Fernández and Krapovickas (1994), have been observed among the landraces (Lavia and Fernández 2004).

The three tetraploid species of section *Rhizomatosae* have their complements mainly constituted by metacentric (m) chromosomes of similar size, with one pair of SAT chromosomes type 3. Among them, *A. pseudovillosa* has a karyotype formula of 40 m, whereas *A. nitida* and both varieties of *A. glabrata* have 38 m + 2 sm (Fernández and Krapovickas 1994; Peñaloza and Valls 2005; Ortiz et al. 2014).

Among the diploid species, only those from the sections Arachis, Caulorrhizae, Heteranthae,

and Rhizomatosae were comprehensively studied. Within species with $2n = 20$ of section Arachis, 15 are characterized by the presence of the A9 chromosomes and have symmetric karyotypes mainly composed of metacentric chromosomes (Fernández and Krapovickas 1994; Lavia 1996, 2000; Peñaloza and Valls 2005; Robledo et al. 2009). The most frequent karyotype formula is 18 m + 2 sm. The remaining species with $2n = 20$ (genomes B, F, and K, after Robledo and Seijo 2010) have symmetric karyotypes, without any small chromosome (Smartt et al. 1978; Smartt and Stalker 1982; Fernández and Krapovickas 1994; Lavia 1996; Peñaloza and Valls 2005). The most frequent karyotype formula is 20 m, but one to four pairs of submetacentric chromosomes were observed in different species (Fernández and Krapovickas 1994; Robledo and Seijo 2010). *Arachis glandulifera* (D genome) is the only species ($2n = 20$) that has an asymmetric karyotype with a karyotype formula of 8 m + 4 sm + 8 st (Stalker 1991; Robledo and Seijo 2008). Diploid species with $2n = 18$ (G genome, after Silvestri et al. 2015) have symmetric karyotypes with all the chromosomes of similar size (Lavia 1996, 1998; Peñaloza et al. 1996). Almost all the species of the section Arachis have only one pair of secondary constrictions localized on the long arms of pair 10. The exceptions are *A. krapovickasii* and *A. gregoryi*, which have two pairs of secondary constrictions (Robledo and Seijo 2008, 2010; Robledo et al. 2009).

The species of section *Heteranthae* have a karyotype formula of 18 m + 2 sm and SAT chromosomes type 2 (*A. dardani*, *A. pusilla*, and *A. interrupta*), or 16 m + 4 sm and with SAT chromosomes type 10 (*A. sylvestris* and *A. giacomettii*) (Silva et al. 2010). The karyotype formula of the *Caulorrhizae* species is 18 m + 2 sm, but *A. pintoi* has a pair of SAT chromosomes type 2, while *A. repens* has type 3 (Pucciariello et al. 2013). The formula of *A. burkartii* of section *Rhizomatosae* is 20 m with SAT chromosomes type 8 (Ortiz et al. 2013).

4.3 Heterochromatin

Chromatin has several classes of proteins complexed with its DNA which are responsible for the major levels of its compaction. Of these levels of chromatin compaction, one of the most conspicuous is the fraction that remains highly condensed throughout the cell cycle, usually defined as constitutive heterochromatin. Although heterochromatin is largely the gene-poor fraction, nowadays it is considered that it plays important roles in the modulation of gene expression, in chromosome structure, and in speciation and evolution of eukaryotes (reviewed in Grewal and Moazed 2003). Accordingly, the analyses of heterochromatin in *Arachis* species revealed extremely different patterns among species and provided many chromosome and genome markers (Seijo et al. 2004; Robledo and Seijo 2008, 2010; Robledo et al. 2009; Custodio et al. 2013; Silvestri et al. 2015).

In the complements of peanut and *A. monticola*, half of the chromosomes (those of the A genome) have centromeric $DAPI^+$ bands (AT-rich), while the remainder (those of the B genome) lack detectable centromeric heterochromatin. The bands of the A complement differ in size, with the most conspicuous being those borne by the A9 pair. All the *A. hypogaea* varieties and *A. monticola* have a similar distribution and amount of heterochromatin, which accounts for about 7% of the karyotype length (Seijo et al. 2004).

Among the diploid species of section Arachis, one of the most variable characters is the amount and distribution of heterochromatin. The A genome species are characterized by having conspicuous $DAPI^+$ centromeric bands in all (or nine) of the chromosome pairs and a total amount of heterochromatin of around 12% of the karyotype length. The size of the bands is around 10% of the chromosome length, except in the A9 and A10 pairs, in which the band size ranged from 25 to 46% and from 14.50 to 20.12%, respectively, among species (Robledo et al. 2009).

The pattern of DAPI$^+$ heterochromatin in the K genome species is characterized by conspicuous centromeric bands in nine chromosome pairs (Robledo and Seijo 2010). The total amount of centromeric heterochromatin per haploid complement is around 12%. Pairs K9 and K10 have the largest bands (16–20%), and the remaining chromosomes have the smallest ones (around 10%) in relation to the chromosome length.

The karyotypes of the F genome species characteristically had small and faint DAPI$^+$ centromeric bands in only seven or eight chromosome pairs (Robledo and Seijo 2010). All the bands had a similar size (around 8% of the chromosome length), and the total amount of heterochromatin per haploid complement was almost half (around 6.5%) of that observed in the species of the A and K genomes.

The B genome species had karyotypes devoid of detectable centromeric bands. However, few species had one small interstitial or distal band in the short arms of pair B3, which usually covers less than 10% of the chromosome length (Robledo and Seijo 2010).

The three $x = 9$ species of section *Arachis* have centromeric DAPI$^+$ bands with the same brightness, position, and size in all chromosome pairs, except *A. palustris*, which lacks these bands in pair G5 (Silvestri et al. 2015).

Aside the species of section *Arachis*, karyotypes with centromeric DAPI$^+$ bands in all (or most) of the chromosome pairs are found in all the sections so far analyzed—Erectoides, Heteranthae, Procumbentes, Triseminatae, Caulorrhizae, and Rhizomatosae (Raina and Mukai 1999; Lavia et al. 2011; Pucciariello et al. 2013; Ortiz et al. unpublished). CMA$^+$/DAPI$^-$ (GC-rich) heterochromatin is restricted to the secondary constrictions of SAT chromosomes in the species of sections Arachis, Caulorrhizae, and Heteranthae studied so far, except *A. dardani* and *A. giacomettii* that present centromeric CMA$^+$ bands in one chromosome pair (pairs 3 and 5, respectively) and *A. pusilla* that has centromeric GC-rich heterochromatin in the centromeric regions of all the chromosomes (Silva et al. 2010).

4.4 rDNA Loci

Localization of the 5S and 18S-26S rRNA genes on the chromosomes by fluorescent in situ hybridization (FISH) was initially applied to a small set of species from different sections of *Arachis* (Raina and Mukai 1999) revealing their usefulness for the characterization of the species. Chromosome mapping of these loci was later used to analyze in detail the karyotypes of all the species included in section *Arachis* (Seijo et al. 2004; Robledo and Seijo 2008; Robledo et al. 2009, 2010; Robledo and Seijo 2010; Lavia et al. 2011; Custodio et al. 2013; Silvestri et al. 2015).

Physical mapping of the rDNA loci in the six botanical varieties of *A. hypogaea* and in *A. monticola* revealed two pairs of 5S and five pairs of 18S–26S rDNA sites. In both species, the 5S loci are proximally located in short arms (pairs A3 and B3), while the 18S–26S rDNA loci are proximally (pairs A2, A10, B3, and B10) or subterminally placed (B7). One 5S locus is syntenic with a 18S–26S site in the pair B3. The high degree of homeology detected between *A. monticola* and *A. hypogaea* strongly evidences that they are very closely related taxa. The mapping of the rDNA loci, together with the heterochromatin analysis, provided the first chromosome map for peanut (Seijo et al. 2004).

Regarding diploid species of section Arachis, those within the A genome have only one interstitial (or rarely proximal) 5S rDNA locus located in the A3 pair. However, the number, size, and chromosomal localization of the 18S–26S rDNA loci vary among the species (Robledo et al. 2009). The number of these gene clusters ranges from two to four pairs, and they present variable size. In general, the largest loci are located in pair A10, those of intermediate size in pair A2, while the smallest and faintest signals (in the cases that the species have more than two loci) in the pairs A7 and A4. According to the pattern of rDNA loci and the heterochromatic bands, the A genome species have been further arranged into three karyotype groups (Robledo et al. 2009): Chiquitano, Pantanal, and La Plata River Basin.

Arachis glandulifera (D genome) has only one 5S rDNA locus subterminally located in long arms of pair D5 and five 18S–26S rDNA loci located in different positions and arms of the D1, D2, D6, D9, and D10 pairs. The markers so far identified were enough to the precise identification of all the chromosome pairs of the karyotype and to the construction of the first wholly resolved idiogram for an *Arachis* species (Robledo and Seijo 2008).

All the other species of section Arachis have one 5S rDNA locus localized in proximal or interstitial position on the short arms of the metacentric pair 3. Exceptions are the species of the K genome that have two additional pairs of loci located in the pairs K8 and K10. The number of 18S–26S rDNA loci ranges from two (in *A. gregoryi* and *A. trinitensis*) to four (in *A. magna* and *A. valida*). Most of them are located in pericentromeric or interstitial position on the long arms. In most species, the largest and brightest 18S–26S rDNA loci correspond to the clusters located in the secondary constrictions of the SAT chromosomes (pair 10), while the remaining ones are small and pale. In some species, one 18S–26S rDNA and one 5S rDNA loci are localized on the same chromosome. In the species of the K genome, these loci co-localize on the long arm of pair K10, while in *A. benensis* and *A. magna*, they co-localize on the short arm of pairs F3 and B3, respectively. In *A. ipaënsis*, one 18S–26S rDNA and one 5S rDNA loci map to pair B3, but on different arms (Robledo and Seijo 2010).

The three species with $x = 9$ have only one 18S–26S rDNA site in the proximal region of the long arm of the G9 pair (SAT chromosomes) and one 5S rDNA site in the short arm of the pair G6 (Silvestri et al. 2015).

4.5 Repetitive Sequences

Plant genomes are composed of single-copy sequences, with one or few copies (gene sequences), and repetitive sequences, with a higher copy number. The latter can be found as dispersed repetitive (transposons) or tandemly repetitive (satellite DNA) sequences (Schmidt and Heslop-Harrison 1998). One of the most important features of the repetitive genome component is its rapid evolution both at the sequence level and genome representation (Schmidt and Heslop-Harrison 1998). For this reason, the analysis of this fraction is a useful tool for the study of evolutionary relationships between plant species (Dechyeva et al. 2003; Navajas- Pérez et al. 2009; Nielen et al. 2010, 2012; Samoluk et al. 2015a). Moreover, many of these sequences when probed onto chromosomes provide conspicuous markers for genome characterization, for the establishment of homeologies and for the construction of chromosome maps (Seijo et al. 2004; Robledo et al. 2009; Robledo and Seijo 2010; Zhang et al. 2016).

4.6 Dispersed Sequences

Studies based on retrotransposons at a genomic scale are scarce in *Arachis* and have been focused on peanut and its wild diploid progenitors (*A. ipaënsis* and *A. duranensis*). To date, Ty3-gypsy (Nielen et al. 2010) and Ty1-copia (Nielen et al. 2012) LTR retrotransposons have been characterized and quantified in these three species. A significant differential representation of Ty3-gypsy retrotransposons, but not of Ty1-copia retrotransposons, was described in the two diploid species. The element, named FIDEL (Fairly long Inter-Dispersed Euchromatic LTR retrotransposon), is more frequent in the A than in the B genome, with copy numbers of about 3000 (± 950, *A. duranensis*), 820 (± 480, *A. ipaënsis*), and 3900 (± 1500, *A. hypogaea*) per haploid genome. Phylogenetic analysis of reverse transcriptase sequences showed the distinct evolution of FIDEL in the diploid species. Fluorescent in situ hybridization revealed a disperse distribution in the euchromatin and absence from centromeres, telomeric, and the nucleolar organizer regions. Distribution of FIDEL onto the chromosomes reflects almost the pattern of GISH using genomic probes of the diploid progenitors onto the chromosomes of peanut (Seijo et al. 2007).

By contrast, the Ty1-copia retrotransposon from the Bianca lineage (named Matita) is a moderate copy number element (Nielen et al. 2012). This element is almost equally represented in the A and B genomes in relatively low copy. FISH experiments showed that Matita is mainly located in the distal regions of the chromosome arms and its chromosome-specific hybridization pattern aided in the identification of some individual chromosomes. By probing BAC libraries with overgos probes from Matita, it was demonstrated that this element is not randomly distributed in the genome, but exhibits a significant tendency of being more abundant near resistance gene homologues than near single-copy genes.

A more recent study comparing 1.26 Mb of homeologous A and B genomes BAC clones evidenced the existence of a diverse group of complete and truncated copies of the LTR retrotransposons fraction that covered more than 40% of the sequences analyzed (Bertioli et al. 2013). BAC-FISH using 27 *A. duranensis* BAC clones as probes gave dispersed and repetitive DNA characteristic signals, predominantly in interstitial regions of the peanut A chromosomes. In general, the sequences of 14 BAC clones revealed that a substantial proportion of the highly repetitive component of the peanut A genome is represented for relatively few LTR retrotransposons and their truncated copies of LTRs.

Non-LTR retroelements are generally less abundant than LTR retroelements in the plant genomes (Noma et al. 1999; Alix and Heslop-Harrison 2004; Hawkins et al. 2006), and *Arachis* genomes are not the exception. The available data reveal that the genome content of these elements is less than 8% in *A. duranensis* (Chen et al. 2016) and approximately 12% in *A. ipaënsis* (Bertioli et al. 2013). However, the genome representation of these retroelements is relatively high when compared to the amount of LINEs present in other plant species (Samoluk et al. 2015b; Bertioli et al. 2016). Recently, the diversity, the chromosome distribution, and the genome representation of a LINE family belonging to the L1 clade were analyzed in six genomes and karyotype groups of section *Arachis* (Samoluk et al. 2015b). The phylogenetic analysis based on the reverse transcriptase of these elements showed that the lineages are distributed independently of the genomes or karyotype groups. FISH experiments revealed a dispersed pattern with hybridization signals mainly located on the euchromatin of interstitial and distal regions of most chromosome arms in all the genome types analyzed in that study. In agreement with the results obtained by Bertioli et al. (2016), the genome abundance of this kind of retroelements was higher in *A. ipaënsis* than in *A. duranensis*.

The sum of available data evidences that retroelements have a dispersed pattern in all the genomes analyzed so far, although with different representation among them. Though some of these elements may be used for the identification of different chromosome complements in hybrids and allopolyploids, their usefulness as cytogenetic markers for individual chromosome identification is limited.

In contrast to the high genome abundance of retroelements, DNA transposons constitute about 5–10% of the genome of *Arachis* species (Bertioli et al. 2016; Chen et al. 2016). However, these elements have been little explored in *Arachis*, and there is no available data on their chromosome distribution. In this sense, there are some reports about a miniature inverted-repeat transposable element, the AhMITE1 element (Shirasawa et al. 2012; Gowda et al. 2010, 2011). This element is present in high copy number in the genomes of *A. hypogaea*, *A. magna*, and *A. monticola*, but not in *A. duranensis* (Shirasawa et al. 2012). In addition, the study of AhMITE1 transpositional activity among different allotetraploid *Arachis* species showed that most of the *A. hypogaea* subsp. *fastigiata* types carry an AhMITE1 insertion at the FST1-linked site, whereas the wild allotetraploid *A. monticola* and *A. hypogaea* subsp. *hypogaea* types missed the AhMITE1 element at that site (Gowda et al. 2011). This finding supports that the *A. hypogaea* subsp. *hypogaea* is closer to the wild allotetraploid than the subsp. *fastigiata* (Paik-Ro et al. 1992; Singh et al. 1993; He and Prakash 2001).

Moreover, it proposed that AhMITE1 transposition could have been of major importance in the origin of *A. hypogaea* subsp. *fastigiata* (Gowda et al. 2011).

4.7 Clustered Sequences

Satellite DNA constitutes a significant portion of eukaryote genomes. It is formed by repetitive units of variable length (140–180 bp or 300–360 bp) tandemly arranged in blocks of up to 100 Mpb (Charlesworth et al. 1994; Schmidt and Heslop-Harrison 1998; Plohl et al. 2008). These sequences usually show particular chromosome locations, being a major component of the centromeric (Hudakova et al. 2001; Gindullis et al. 2001; Urdampilleta et al. 2009), telomeric (Pich et al. 1996; Macas et al. 2000), and less frequently, interstitial heterochromatin (Mukai et al. 1992). Therefore, they have become a useful tool to study the karyotype evolution in different groups of species (Lanfredi et al. 2001; Slamovits et al. 2001).

The analysis of a satellite sequence named ATR-2 in seven diploid species ($x = 10$) representative of different genomes and karyotype groups revealed an infraspecific and interspecific conservation of these sequences, with a low spreading of new monomeric variants in the analyzed species (Samoluk et al. unpublished). However, the quantitative analyses revealed differences in the abundance of this satellite DNA among them, according to the predictions of the "library hypothesis" (Fry and Salser 1977). FISH analyses revealed that ATR-2 is exclusively distributed at the $DAPI^+$ centromeric heterochromatin; however, it may not be the only sequence that conforms this genomic fraction. Despite the sequence conservation of ATR-2, the variable representation of this satDNA suggests that it was actively involved in the remodeling the heterochromatic patterns of the diploid *Arachis* species (Samoluk et al. unpublished).

Another major repetitive DNA sequences were cloned and analyzed from *A. hypogaea* Cot-1 DNA in order to identify new genome- and chromosome-specific markers (Zhang et al. 2012). In particular, a satellite DNA sequence of 115 bp was found mainly distributed in pericentromeric regions on most of the B genome chromosomes of the allotetraploid (Zhang et al. 2012). Recently, it was developed an FISH-based karyotyping system using a set of new and previously reported chromosome markers, which allowed the identification of almost all chromosomes and the construction of karyotypes in cultivated peanut and its two putative progenitors (Zhang et al. 2016).

4.8 Genome Size

Genome size is a useful descriptor for characterization of plant genetic resources (Ozias Akins and Jarret 1994; Rayburn et al. 1997; Hendrix and Stewart 2005; Loureiro et al. 2007). Until recently, nuclear DNA contents were studied in a very limited number of *Arachis* species, and the available genome size estimations were controversial. Most of these determinations were made by the Feulgen densitometry method (Dhillon et al. 1980; Resslar et al. 1981; Singh et al. 1996; Lavia and Fernández 2008). However, measurements in *A. duranensis* and *A. hypogaea* by flow cytometry indicated that the data obtained by Feulgen densitometry [except in Dhillon et al. (1980) for *A. hypogaea*] overestimated the genome size of *Arachis* species by twofold (Temsch and Greilhuber 2000, 2001).

In a recent study, 26 diploid species of the section *Arachis* were analyzed by flow cytometry, and the 2C values ranged from 2.55 to 3.22 pg (Samoluk et al. 2015b). The annual species belonging to different genomes (Robledo et al. 2009; Robledo and Seijo 2010) tend to have different genome sizes. However, the 2C values of the perennial species (all with A genome) are distributed almost continuously along the whole range of genome sizes (2.55–3.22 pg) of the annual species. The comparison of 2C values with karyotype parameters suggests that changes in DNA content have been proportionally distributed among the chromosome arms, and that the heterochromatic fraction is not directly involved in those changes. Within the A genome,

the annual species has lower DNA content than the perennial ones, which is in accordance with the predictions of the nucleotype hypothesis (Bennet 1982). However, the lack of significant relationships with geoclimatic variables suggests that there are many intrinsic factors determining particular ecological roles of the DNA content in different lineages of section Arachis. A critical analysis of the DNA content of other species of *Arachis* is still needed to address the direction of the genome change during the evolution of the genus as a whole.

Measurements of the genome size of the AABB species of *Arachis* showed that they are among the few allopolyploids in which their genome size are about the sum of those of their diploid progenitors (Samoluk et al. 2015b), like in tobacco (Leitch and Bennett 2004), *Hordeum* (Jakob et al. 2004), and AD *Gossypium* species (Wendel et al. 2002). The genome sizes estimated for *A. monticola* (5.70 pg) and for the cultivated peanut (5.60 pg) were *on par* with the expected 2C value estimated from the sum of the genome sizes of their parental species (*A. duranensis*, 2C = 2.55 pg, and *A. ipaënsis*, 2C = 3.19 pg) (Samoluk et al. 2015b). The constancy in the Cx values suggests that the hybridization and chromosome doubling events that occurred during the origin of the cultivated peanut have not resulted in significant changes in genome size.

4.9 Usefulness of Chromosome Markers

The use of chromosome markers in *Arachis* species came to complete and extends a large number of the taxonomic, classical cytogenetic, cross hybridization, molecular marker, and geographical studies. Here we describe two cases in which the use of these chromosome markers contribute to shed light on long debates among researchers: one is about the origin of peanut and the other deals with the genome characterization of the *Arachis* species.

4.9.1 Inferences on Peanut Origin

The origin of peanut has been a matter of study for several decades and has long been assessed from different points of view. The identification of one pair of small chromosomes (A9 pair after Robledo et al. 2009), and one pair of SAT chromosomes (A10 pair after Fernández and Krapovickas 1994), instead of four chromosomes of each type led to the proposal that the peanut is an allotetraploid species with $2n = 4x = 40$ (Husted 1933, 1936), and with an AABB genome constitution (Smartt et al. 1978). This has been confirmed by studies on interspecific hybridization among the cultivated peanut and different wild diploid species (Smartt and Gregory 1967; Stalker and Wynne 1979; Singh 1986), and by modern cytogenetic techniques (Seijo et al. 2004, 2007).

However, the diploid species that were involved in the origin of cultivated peanut were until recently under debate. Before the development of chromosome markers, more than eight wild diploid species having different genome types were considered involved in the origin of peanut (reviewed in Singh and Smartt 1998; Seijo et al. 2007; Grabiele et al. 2012). Studies based on molecular markers showed that several species of the A genome could be considered as the most probable ancestor of peanut. Restriction fragment length polymorphism (RFLP) revealed *A. duranensis* as the most probable candidate (Kochert et al. 1991, 1996), whereas randomly amplified polymorphic DNA (RAPD) and inter-simple sequence repeat (ISSR) analyses showed *A. villosa* as the best candidate (Raina et al. 2001). On the other hand, PCR amplified fragment length polymorphism (AFLP) data have shown that at least three diploid species of the A genome were closely related to the cultigen (Milla et al. 2005). Similarly, microsatellite markers have revealed that, although *A. duranensis* is the most closely related to the cultigen, a small group of other species having the A genome could also be genome donors of peanut (Moretzsohn et al. 2004).

Classical cytogenetics aided in the identification of the diploid progenitors of peanuts. However, only when chromosome banding and molecular cytogenetic techniques were applied massively to *Arachis* species by determining the patterns of heterochromatin and the number (Raina and Mukai 1999) and position (Seijo et al. 2004; Robledo et al. 2009; Robledo and Seijo 2010; Custodio et al. 2013) of the rRNA genes by FISH, the wild diploid progenitors were more precisely identified. The analysis of the rDNA loci distribution showed that the species of the A genome included in the La Plata River Basin group are the most related to the A genome of the tetraploids, and that *A. ipaënsis* is the most probable B genome donor. Subsequently, double GISH experiments using genomic DNA of the diploid *Arachis* species ($2n = 2x = 20$) identified *A. duranensis* as the A genome donor of *A. hypogaea* (Seijo et al. 2007).

Arachis monticola is the only wild allotetraploid within section *Arachis*, and it is currently known only from three very close localities in NW of Argentina. In all the dendrograms constructed using molecular markers, the cultivated peanut and *A. monticola* group together, with very low or no genetic distance (Gimenez et al. 2002; Grabiele et al. 2012; Moretzohn et al. 2013). They are known to be interfertile, with no apparent sterility in the F_1 hybrids (Kirti et al. 1983). Therefore, *A. monticola* is regarded by different authors either as the direct progenitor of the peanut or as an introgressive derivative between the peanut and wild species (see Stalker and Moss 1987; Grabiele et al. 2012). However, if *A. hypogaea* (AABB) could cross with any diploid species of the A or B genomes through reduced gametes (n), the result would be an infertile or very poorly fertile triploid with AAB or ABB genome constitutions, respectively. Alternatively, if crosses between *A. hypogaea* and diploid species of the A or B genomes had occurred via unreduced gametes ($2n$) of the wild species, then the hybrids would have AAAB or ABBB genome constitutions, respectively. Neither of these scenarios would have given rise to an allotetraploid with the genome constitution compatible with that of *A. monticola* (AABB) (Seijo et al. 2004, 2007; Grabiele et al. 2012). This rationale suggested that the hypothesis that considers *A. monticola* as an introgressive derivative has to be discarded in the light of the molecular cytogenetic data, since the wild tetraploid has exactly the same chromosome complements to that of *A. hypogaea.*

The fact that the amphidiploid that resulted from the artificial resynthesis from *A. ipaënsis* and *A. duranensis* (Fávero et al. 2006) is morphologically very similar to *A. monticola*, and that it can hybridize with all the varieties of the cultigen producing fertile offspring, supports the hypothesis that considers *A. monticola* as the direct progenitor of *A. hypogaea*. Moreover, the ability to persist in natural populations (unlike the cultivated peanut) and the wild type structure of its fruits (wherein each seed has its own shell separated by an isthmus) support the maintenance of *A. monticola* as a separate taxonomic species.

The identical patterns of molecular cytogenetic markers and genomic hybridization (GISH) detected in all the botanical varieties of the cultigen suggests that the same wild species participated in their origin. Moreover, this finding implies that all the presently known varieties of peanut arose from a single, unique allotetraploid plant population (Seijo et al. 2004; Seijo et al. 2007). The common ancestry of all infraspecific taxa of *A. hypogaea* is supported by the low genetic variability so far detected with most molecular markers in the cultivated peanut (Halward et al. 1991; Kochert et al. 1996; Herselman 2003; Grabiele et al. 2012; Moretzohn et al. 2013).

In this scenario, after the origin of the wild allotetraploid (which probably had larger seeds than any of the progenitors as a result of the gigas effects in polyploids, like in *A. monticola*), *A. hypogaea* may have arisen through domestication. Therefore, the large morphological, ecological, phenological, and chemical variability present in the many landraces of peanut (Krapovickas et al. 2009, 2013) would have mainly resulted from particular selective

pressures undergone in different agroecological environments (Krapovickas and Gregory 1994; Grabiele et al. 2012).

4.9.2 Origin of Other Tetraploid *Arachis* Species

Section *Rhizomatosae* is currently defined exclusively on morphological features, mainly because all (four) the species have rhizomes and can be asexually propagated. Among this group of taxa, *A. burkartii* is the only diploid with $2n = 2x = 20$, while the others, *A. pseudovillosa, A. nitida*, and *A. glabrata*, are tetraploids with $2n = 4x = 40$ (Gregory et al. 1973; Fernández and Krapovickas 1994; Peñaloza and Valls 2005). Interestingly, the tetraploid species are distributed from the Mato Groso State in Brazil to the North of Argentina, while the diploid species live from the North of Argentina to the North of Uruguay. Thus, the species with different ploidy level of section *Rhizomatosae* only overlaps in a very narrow stretch in the NE of Argentina (Krapovickas and Gregory 1994; Valls and Simpson 2005).

If this section were monophyletic, it would be expected that *A. burkartii* was one of the genome donors of the rhizomatous tetraploids. However, classical cytogenetic and molecular data argued the monophyly of the section and the nature of the polyploids. The meiotic analysis of *A. nitida* revealed that the chromosomes are arranged in 20II in most (65%) of the cells, although up to four multivalents (trivalents and quadrivalents) in low frequencies were observed (Ortiz et al. 2011). Moreover, it has one 5S rDNA locus, two 18S–26S rDNA loci, and a similar pattern of centromeric heterochromatic bands in the four chromosome sets. However, the presence of one 18S–26S rDNA sites in only two of the four chromosomes of pair 2 of *A. nitida* suggests that it has two different chromosome sets. Thus, *A. nitida* may be considered a segmental allopolyploid (Ortiz et al. unpublished).

Controversial hypotheses have been put forward for the nature of *A. glabrata*. Some authors proposed that this species is a true autopolyploid (Singh and Simpson 1994; Ortiz et al. 2011), while others suggested that it may be an allopolyploid with the EERR genome constitution (Gregory and Gregory 1979; Bechara et al. 2010). Meiotic analyses of this species showed that the frequency of bivalents differs significatively (from 20 to 81%) among accessions (Ortiz et al. 2011). Also, the number of trivalents (from 1 to 3III) and quadrivalents (from 1 to 8IV) was very variable among them. However, based on the fact that the four chromosome sets have $DAPI^+$ centromeric bands in all the chromosomes and a similar pattern of the 5S and 18S–26S rDNA loci, together with the detection of up to eight quadrivalents in meiotic cells, it was suggested that this species may be either a true autopolyploids or, less probably, a segmental allopolyploid with different degrees of diploidization (Ortiz et al. unpublished).

Concerning *A. pseudovillosa*, since meiotic behavior studies could not be performed yet, the polyploidy nature is still under study (Ortiz et al. unpublished). In this sense, the presence of one 5S and one 18-26S rDNA sites, and the similar distribution pattern of CMA–DAPI bands in the four chromosome sets suggest that *A. pseudovillosa* may be an autopolyploid. However, the presence of an extra 18S–26S rDNA locus in only one chromosome set may have arisen de novo by transposition (or other genomic mechanism) after polyploidization, or may be interpreted as evidence for a segmental allopolyploid origin by hybridization between two closely related species which differ in the numbers of this rDNA cluster genes (Ortiz et al. unpublished).

Molecular markers strongly support the conclusion obtained from the cytogenetic data about the relationship among the rhizomatous species. RAPD (Nobile et al. 2004) and SSR (Angelici et al. 2008) analyses that included the four species of section *Rhizomatosae* showed the clustering of the tetraploid species in a single group, distant from *A. burkartii*. Further, the AFLP analysis including representatives of seven different sections revealed a close association of *A. glabrata* with *A. major* and *A. paraguariensis* (sect. *Erectoides*), while *A. burkartii* was associated with the two species of the section

Caulorrhizae (Gimenez et al. 2002). In addition, molecular phylogenies based on chloroplast and nuclear DNA sequences (Bechara et al. 2010; Friend et al. 2010), which only included *A. burkartii* and *A. glabrata*, have shown that the diploid taxon is found in an individual and isolated clade, while *A. glabrata* grouped in a distant clade with members of sections *Erectoides* and *Procumbentes.*

Regarding the genomic constitution of rhizomatous tetraploid species, the cytogenetic evidences suggest that the three species may have at least one common diploid ancestor (Ortiz et al. unpublished). In this sense, all the species of the section *Rhizomatosae* (2*x* and 4*x*) have been traditionally assigned to the R genome (Smartt and Stalker 1982) assuming that the section was monophyletic and the diploid rhizomatous species, *A. burkartii,* was the natural ancestor of the tetraploids. However, the fact that *A. burkartii* has a karyotype formula of 20 m with a SAT chromosome type 8, lacks $DAPI^+$ heterochromatic bands, and presents four pairs of 18S–26S rDNA loci (located on three chromosome pairs) and one pair of interstitial 5S rDNA sites (that co-localized with of 18S–26S rDNA loci on the smallest chromosome pair) evidenced that this diploid species is unlikely the genome donor of the rhizomatous tetraploids. Instead, the complements of these tetraploids showed high homology with those of the *Erectoides* and *Procumbentes* sections. Thus, the tetraploid species should be excluded from the R-genome, which might remain exclusively for *A. burkartii* (Ortiz et al. unpublished).

The aforementioned cytogenetic data do not support the origin of rhizomatous tetraploids from *A. burkartii* and suggest that the section *Rhizomatosae* is not monophyletic (Nóbile et al. 2004; Angelici et al. 2008; Bechara et al. 2010; Friend et al. 2010; Ortiz et al. unpublished).

4.9.3 Genome Arrangement of Section *Arachis*

Diploid species of section *Arachis* with 2*n* = 20 and symmetric karyotypes have been traditionally arranged first into two different genome groups (A and non-A genomes) on the basis of the presence of the small chromosomes A9 (first observed in *A. hypogaea*) in their karyotypes. These two groups of species show strong reproductive isolation manifested by low hybrid production, and low chromosome pairing and low pollen stainability in the hybrids (Gregory and Gregory 1979; Stalker et al. 1991; Krapovickas and Gregory 1994; Tallury et al. 2005). The only species with 2*n* = 20 and asymmetric karyotype (*A. glandulifera*) was assigned to a different genome, the D genome (Stalker 1991).

The development of chromosome markers by molecular cytogenetics revealed a high degree of homogeneity in the karyotypes among the species with A chromosomes. However, variation in number and positions of $DAPI^+$ bands and major 18S–26S rDNA sites among species was used to establish three subgroups of karyotype homeologies (Robledo et al. 2009) considering the fact that closeness of taxa is usually correlated with the similarity of their heterochromatin and rDNA FISH patterns (Hizume et al. 2002; Liu et al. 2003). Since the groups that resulted from the homeology analysis included species that tend to be more closely distributed geographically than those belonging to different groups, they were named using a geographical reference (Robledo et al. 2009). The Chiquitano group comprised the species (*A. cardenasii*, *A. herzogii*, and *A. kempff-mercadoi*) that grow in the southern and western portion of the Chiquitanía biogeographic region in Santa Cruz Department of Bolivia. The Pantanal group includes the species (*A. diogoi*, *A. kuhlmannii*, *A. helodes*, *A. simpsonii*, and *A. stenosperma*) which are distributed in the Pantanal biogeographic region in western Brazil, northern Paraguay, and eastern Bolivia. This group may also include *A. linearifolia*. The La Plata River Basin group corresponds to the species (*A. duranensis*, *A. schininii*, *A. correntina, A. villosa*, and probably *A. microsperma*) that are distributed along the La Plata River Basin (except the region comprising the upper stream of the Paraguay River in the Pantanal).

Most molecular marker studies in a large set of A genome species support the Pantanal group (Kochert et al. 1991; Raina et al. 2001; Milla

et al. 2005; Moretzon et al. 2013). The Chiquitano group is least represented in molecular analyses, but whenever *A. herzogii* and *A. kempff-mercadoi* have been included, they always clustered together (Milla et al. 2005; Tallury et al. 2005). In a recent microsatellite-based phylogenetic analysis, *A. cardenasii* and *A. kempff-mercadoi* were grouped together with the species that belong to the Pantanal group (Moretzohn et al. 2013). However, only two accessions of *A. cardenasii* and none of *A. herzogii* were included in that study. Species belonging to La Plata River Basin group generally clustered together with a few exceptions (Milla et al. 2005; Moretzohn et al. 2013).

The non-A genome species have been segregated into three genomes (B, F, and K) based on the different patterns of chromosome markers (Robledo and Seijo 2010). The K genome includes *A. batizocoi*, *A. cruziana*, and *A. krapovickasii.* These species are characterized by having conspicuous $DAPI^+$ heterochromatic bands in all chromosome pairs except in K7 and three 5S rDNA loci on K2, K4, and K10 pairs. The F genome comprises *A. benensis* and *A. trinitensis*, and their karyotypes have small and faint bands in seven or eight chromosome pairs and only one 5S rDNA locus on F3 pair. The B genome species (after Robledo and Seijo 2010) have karyotypes without pericentromeric $DAPI^+$ heterochromatin and includes *A. ipaënsis*, *A. gregoryi*, *A. magna, A. valida*, and *A. williamsii* (Robledo and Seijo 2010). This group showed the highest homeology with the B genome of *A. hypogaea*. *Arachis glandulifera* has asymmetric karyotype (Stalker 1991; Fernández and Krapovickas 1994), a unique pattern of heterochromatin $DAPI^+$ distribution and 5 pairs of 18S–26S rDNA loci (Robledo and Seijo 2008). On this basis, this species was confirmed as having the D genome as proposed earlier by Stalker (1991).

The proposed genome arrangement is supported by species crossability, pollen stainability, morphological characters, and geographical distribution of the species (Smartt et al. 1978; Stalker 1991; Krapovickas and Gregory 1994; Tallury et al. 2005; Burrow et al. 2009). The position of the D, F, and K genomes with respect to the A and B genomes is controversial. The analysis of chloroplast sequences revealed that these genomes are closely related to the B genome, but the NTS of the 5S rDNA genes and AFLP markers showed that the F genome is indeed close to the B genome, while the D and K are sister groups of the A genome (Tallury et al. 2005; Grabiele et al. 2012). The phylogenetic analysis based on DNA sequence information of three single-copy gene introns was consistent with the current genome classification, since clades contained species with the same genome types (Moretzshon et al. 2013). In this analysis, the species with D, F, and K genome species are close to the A genome species, but the microsatellite analysis done in the same report showed that those genomes types are closer to the B genome than to the A genome. In a more recent analysis using nine intron sequences and GISH, it was shown that the K genome is closer to the B genome than to the A genome of *A. hypogaea* (Leal-Bertioli et al. 2015). The extant inconsistency of available data showed that the D, F, and K genomes are in an intermediate position between the A and the B genomes.

Geographically, the species included within each genome tend to be co-distributed. The species with the F genome are restricted to the lowland savannas of Beni department in Bolivia, while those with the K genome are distributed in the NW of the Chacoan Boreal region. The species with the B genome are more widely distributed in semi-deciduous forests and savannas of the cerrado associated with the Chiquitano planalto and west Pantanal. *Arachis ipaënsis*, known from only one population, was collected from the top of the sand banks of streams in an ecotone between the tucumano-oranence deciduous forest and the chacoan xerophytic forest (Robledo and Seijo 2010).

More recently, the three species with $x = 9$ chromosomes of section *Arachis* were analyzed with the same chromosome markers that were mapped in the $x = 10$ species. These analyses showed that $x = 9$ species present a similar pattern of $DAPI^+$ heterochromatin to that observed in the species assigned to A and K genomes. However, it differs from the former by lack of the

A9 chromosomes and the presence of only one pair of 18S–26S rDNA sites. From the latter, it differs by the lower number of 5S rDNA sites and the lower number of submetacentric chromosomes. Based on all these karyotypic differences together with the reproductive isolation from any other species of section *Arachis*, a new genome type (G genome) was proposed for the $x = 9$ species (Silvestri et al. 2015). Molecular datasets have revealed that the three $x = 9$ species form a compact clade, but different from any other group in section *Arachis* (Bechara et al. 2010; Friend et al. 2010; Moretzshon et al. 2013). Moreover, their relationship with other species of the section is still unclear. In this sense, analysis of microsatellites (Moretzshon et al. 2004, 2013), single-copy gene sequences (Moretzshon et al. 2013), trnT-F cpDNA marker (Tallury et al. 2005), and ITS and 5.8S of the nuclear rDNA (Bechara et al. 2010) suggest that these taxa are more closely related to some non-A genome species, whereas RAPD (Creste et al. 2005) and AFLP (Milla et al. 2005) analyses suggest that they are genetically more similar to A genome species.

4.10 Conclusion

In conclusion, in spite of the homogeneity in the chromosome morphology of the species of section *Arachis*, the use of chromosome markers revealed six different karyotype organizations, which agree with the different degrees of reproductive isolation. Whereby, based on these different organizations, the species of section Arachis are currently arranged in six different genomes (A, B, D, F, G, and K).

References

Alix K, Heslop-Harrison JS (2004) The diversity of retroelements in diploid and allotetraploid *Brassica* species. Plant Mol Biol 54:895–909

Angelici C, Hoshino AA, Nóbile PM et al (2008) Genetic diversity in Rhizomatosae section of the genus *Arachis* (Fabaceae) based on microsatellite markers. Genet Mol Biol 31:79–88

Bechara MD, Moretzsohn MDC, Palmieri DA et al (2010) Phylogenetic relationships in genus *Arachis* based on ITS and 5.8S rDNA sequences. BMC Plant Biol 10:255. doi:10.1186/1471222910255

Bennet MD (1982) Nucleotypic basis of the spatial ordering of chromosomes in eukaryotes and the implications of the order for genome evolution and phenotypic variation. In: Dover GA, Flavel RB (eds) Genome evolution. Academic Press, New York, pp 239–261

Bertioli DJ, Cannon SB, Froenicke L et al (2016) The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. Nat Genet 48:438–446

Bertioli DJ, Vidigal BS, Nielen S et al (2013) The repetitive component of the A genome of peanut (*Arachis hypogaea*) and its role in remodelling intergenic sequence space since its evolutionary divergence from the B genome. Ann Bot 112:545–559

Burow MD, Simpson CE, Faries MW et al (2009) Molecular biogeographic study of recently described B- and A-genome *Arachis* species, also providing new insights into the origins of cultivated peanut. Genome 52:107–119

Charlesworth B, Sniegowski P, Stephan W (1994) The evolutionary dynamics of repetitive DNA in eukaryotes. Nature 371:215–220

Chen X, Li H, Pandey MK et al (2016) Draft genome of the peanut A-genome progenitor (*Arachis duranensis*) provides insights into geocarpy, oil biosynthesis, and allergens. PNAS 113:6785–6790

Creste S, Tsai SM, Valls JFM et al (2005) Genetic characterization of Brazilian annual *Arachis* species from sections Arachis and Heteranthae using RAPD markers. Genet Resour Crop Evol 52:1079–1086

Custódio AR, Seijo G, Valls JFM (2013) Characterization of Brazilian accessions of wild *Arachis* species of section Arachis (Fabaceae) using heterochromatin detection and fluorescence in situ hybridization (FISH). Genet Mol Biol 36:364–370

Dechyeva D, Gindullis F, Schmidt T (2003) Divergence of satellite DNA and interspersion of dispersed repeats in the genome of the wild beet *B. procumbens*. Chrom Res 11:3–21

Dhillon SS, Rake AV, Miksche JP (1980) Reassociation kinetics and cytophotometric characterisation of peanut (*Arachis hypogaea* L.) DNA. Plant Physiol 65:1121–1127

Fávero AP, Simpson CE, Valls JFM et al (2006) Study of the evolution of cultivated peanut through crossability studies among *Arachis ipaensis*, *A. duranensis* and *A. hypogaea*. Crop Sci 46:1546–1552

Fernández A, Krapovickas A (1994) Cromosomas y evolución en *Arachis* (Leguminosae). Bonplandia 8:187–220

Friend SA, Quandt D, Tallury SP et al (2010) Species, genomes, and section relationships in the genus *Arachis* (Fabaceae): a molecular phylogeny. Plant Syst Evol 290:185–199

Fry K, Salser W (1977) Nucleotide sequences of HS-a satellite DNA from kangaroo rat *Dipodomys ordii* and characterization of similar sequences in other rodents. Cell 12:1069–1084

Gimenes MA, Lopes Romero C, Galgaro ML et al (2002) RFLP analysis of genetic variation in species of section Arachis, genus *Arachis* (Leguminosae). Euphytica 123:421–429

Gindullis F, Desel C, Galasso I et al (2001) The large-scale organization of the centromeric region in beta species. Genome Res 11:253–265

Gowda MVC, Bhat RS, Motagi BN et al (2010) Association of high-frequency origin of late leaf spot resistant mutants with *AhMITE1* transposition in peanut. Plant Breed 129:567–569

Gowda MVC, Bhat RS, Sujay V et al (2011) Characterization of *AhMITE1* transposition and its association with the mutational and evolutionary origin of botanical types in peanut (*Arachis* spp.). Plant Syst Evol 291:153–158

Grabiele M, Chalup L, Robledo G et al (2012) Genetic and geographic origin of domesticated peanut as evidenced by 5S rDNA and chloroplast DNA sequences Plant Syst Evol 298:1151–1165

Gregory WC, Gregory MP (1979) Exotic germplasm of *Arachis* L. interspecific hybrids. J Hered 70:185–193

Gregory WC, Gregory MP, Krapovickas A et al (1973) Structure and genetic resources of peanuts. In: Wilson CT (ed) Peanut culture and uses, Amer. Peanut Res Educ, Soc, Stillwater, pp 47–133

Grewal SIS, Moazed D (2003) Heterochromatin and Epigenetic Control of Gene Expression. Science 301:798–802

Halward TM, Stalker HT, Larue EA et al (1991) Genetic variation detectable with molecular markers among unadapted germplasm resources of cultivated peanut and wild species. Genome 34:1013–1020

Hawkins JS, Kim H, Nason JD et al. (2006) Differential lineagespecific amplification of transposable elements is responsible for genome size variation in *Gossypium*. Genome Res 16:1252–1261

He G, Pakrash C (2001) Evaluation of genetic relationships among botanical varieties of cultivated peanut (Arachis hypogaea L.) using AFLP markers. Genet Resour Crop Evol 4:347–352

Hendrix B, Stewart JM (2005) Estimation of the nuclear DNA content of *Gossypium* species. Ann Bot 95:789–797

Herselman L (2003) Genetic variation among Southern African cultivated peanut (*Arachis hypogaea* L.) genotypes as revealed by AFLP analysis. Euphytica 133:319–327

Hizume M, Shibata F, Matsusaki Y et al (2002) Chromosome identification and comparative karyotypic analyses of four *Pinus* species. Theor Appl Genet 105:491–497

Hudakova S, Michalek W, Presting GG et al (2001) Sequence organization of barley centromeres. Nucleic Acids Res 29:5029–5035

Husted L (1933) Cytological studies on the peanut, *Arachis*. I. Chromosome number and morphology. Cytologia (Tokyo) 5:109–117

Husted L (1936) Cytological studies on the peanut, *Arachis*. II. Chromosome number and morphology and behavior, and their application to the problem of the origin of the cultivated forms. Cytologia (Tokyo) 7:396–423

Jakob SS, Meister A, Blattner FR (2004) The considerable genome size variation of *Hordeum* species (Poaceae) is linked to phylogeny, life form, ecology, and speciation rates. Mol Biol Evol 21:860–869

Kirti PB, Bharathi M, Murty UR et al (1983) Chromosome morphology in three diploid species of *Arachis* and its bearing on the genomes of groundnut (*Arachis hypogaea* L.). Cytologia 48:139–151

Kochert G, Branch WD, Simpson CE (1991) RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. Theor Appl Genet 81:565–570

Kochert G, Stalker HT, Gimenes MA et al (1996) RFLP and cytogenetic evidence on the origen and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). Am J Bot 83:1282–1291

Krapovickas A, Gregory WC (1994) Taxonomía del género *Arachis* (Leguminosae). Bonplandia 8:1–186

Krapovickas A, Vanni RO, Pietrarelli JR et al (2009) Las razas de maní de Bolivia. Bonplandia 18:95–189

Krapovickas A, Vanni RO, Pietrarelli JR et al (2013) Las razas de maní de Perú. Bonplandia 22:19–90

Lanfredi M, Congiu L, Garrido-Ramos MA et al (2001) Chromosomal location and evolution of a satellite DNA familyin seven sturgeon species. Chromosome Res 9:47–52

Lavia GI (1996) Estudios cromosómicos en *Arachis* (Leguminosae). Bonplandia 9:111–120

Lavia GI (1998) Karyotypes of *Arachis palustris* and *A. praecox* (Section Arachis), Two Species with Basic Chromosome Number x = 9. Cytologia (Tokyo) 63:177–181

Lavia GI (2000) Chromosomes studies of wild *Arachis* (Leguminosae). Caryologia 53:177–181

Lavia GI, Fernández A (2004) Karyotypic studies in *Arachis hypogaea* L. varieties. Caryologia 57:353–359

Lavia GI, Fernández A (2008) Genome size in wild and cultivated peanut germplasm. Plant Syst Evol 272:1–10

Lavia GI, Ortiz AM, Robledo G et al (2011) Origin of triploid *Arachis pintoi* (Leguminosae) by autopolyploidy evidenced by FISH and meiotic behavior. Ann Bot 108:103–111

Leal-Bertioli SCM, Santos SP, Dantas KM et al (2015) *Arachis batizocoi*: a study of its relationship to cultivated peanut (*A. hypogaea*) and its potential for introgression of wild genes into the peanut crop using induced allotetraploids. Ann Bot 115:237–249

Leitch IJ, Bennett MD (2004) Genome downsizing in polyploidy plants. Biol J Linn Soc 82:651–663

Liu ZL, Zhang D, Hong DY et al (2003) Chromosomal localization of 5S and 18S-5.8S-25S ribosomal DNA

sites in five Asian pines using Fluorescence in situ hybridization. Theor Appl Genet 106:198–204
Loureiro J, Rodriguez E, Costa A et al (2007) Nuclear DNA content estimations in wild olive (*Olea europaea* L. ssp. *europaea* var. *sylvestris* Brot.) and Portuguese cultivars of *O. europaea* using flow cytometry. Genet Resour Crop Evol 54:21–25
Macas J, Požárková D, Navrátilová A et al (2000) Two new families of tandem repeats isolated from genus Vicia using genomic self-priming PCR. Mol Gen Genet 263:741–751
Milla R, Isleib TG, Stalker HT (2005) Taxonomic relationships among Arachis sect. Arachis species as revealed by AFLP markers. Genome 48:1–11
Moretzhon MC, Gouvea EG, Inglis PW et al (2013) A study of the relationships of cultivated peanut (*Arachis hypogaea*) and its most closely related wild species using intron sequences and microsatellite markers. Ann Bot 111:113–126
Moretzsohn MDC, Hopkins MS, Mitchell SE et al (2004) Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on the analysis of hypervariable regions of the genome. BMC Plant Biol. doi:10.1186/14712229411
Mukai Y, Friebe B, Gill BS (1992) Comparison of C-banding patterns and in situ hybridization sites using highly repetitive and total genomic rye DNA probes of 'Imperial' rye chromosomes added to 'Chinese Spring' wheat. Jpn J Genet 67:71–83
Navajas-Pérez R, Del Bosque MQ, Garrido-Ramos MA (2009) Effect of location, organization, and repeat-copy number in satellite-DNA evolution. Mol Gen Genomics 282:395–406
Nielen S, Campos-Fonseca F, Leal-Bertioli SCM et al (2010) FIDEL-a retrovirus-like retrotransposon and its distinct evolutionary histories in the A- and B-genome components of cultivated peanut. Chromosome Res 18:227–246
Nielen S, Vidigal BS, Leal-Bertioli SCM et al (2012) Matita, a new retroelement from peanut: characterization and evolutionary context in the light of the *Arachis* A-B genome divergence. Mol Genet Genomics 287:21–38
Nóbile PM, Gimenes MA, Valls JFM et al (2004) Genetic variation within and among species of genus *Arachis,* section *Rhizomatosae*. Genet Resour Crop Evol 51:299–307
Noma K, Ohtsubo E, Ohtsubo H (1999) Non-LTR retrotransposons (LINEs) as ubiquitous components of plant genomes. Mol Gen Genet 261:71–79
Ortiz AM, Seijo JG, Fernández A et al (2011) Meiotic behavior and pollen viability of tetraploid *Arachis glabrata* and *A. nitida* species (Section Rhizomatosae, Leguminosae): implications concerning their polyploidy nature and seed set production. Plant Syst Evol 292:73–83
Ortiz AM, Seijo G, Lavia GI (2014) *Arachis.* In: Marhold, K (ed), IAPT/IOPB chromosome data 17. Taxon 63:1153
Ortiz AM, Silvestri MC, Lavia GI (2013) Karyotypic studies in wild species of *Arachis* (Leguminosae) belonging to sections Erectoides, Procumbentes and Rhizomatosae. Bol Soc Argent Bot 48:295–300
Ozias-Akins P, Jarret RL (1994) Nuclear DNA content and ploidy levels in the genus *Ipomoea*. J Am Soc Hortic Sci 119:110–115
Paik-Ro OG, Smith RL, Knauft DA (1992) Restriction fragment length polymorphism evaluation of six peanut species within the Arachis section. Theor Appl Genet 84:201–208
Peñaloza APS, Pozzobon MT, Valls JFM (1996) Cytogenetic findings in wild species of *Arachis* (Leguminosae). In: Sociedade Brasileira de Genética (ed) 42 Congresso Nacional de Genetica, Caxambu, September, 1996, Revista Brasilera de Genetica 19:129
Peñaloza APS, Valls JFM (2005) Chromosome number and satellited chromosome morphology of eleven species of *Arachis* (leguminosae). Bonplandia 14:1–8
Pich U, Fritsch R, Schubert I (1996) Closely related *Allium* species (Alliaceae) share a very similar satellite sequence. Plant Syst Evol 202:255–264
Plohl M, Luchetti A, Meštrović N et al (2008) Satellite DNAs between selfishness and functionality: Structure, genomics and evolution of tandem repeats in centromeric heterochromatin. Gene 409:72–82
Pucciariello O, Ortiz AM, Fernández A et al (2013) Análisis cromosómico del híbrido *Arachis pintoi* x *A. repens* (Leguminosae) mediante citogenética clásica y molecular. Bol Soc Argent Bot 48:111–119
Raina SN, Mukai Y (1999) Detection of a variable number of 18S–5. 8S–26S and 5S ribosomal DNA loci by fluorescent in situ hybridization in diploid and tetraploid Arachis species. Genome 59:52–59
Raina SN, Rani V, Kojima T et al (2001) RAPD and ISSR fingerprints as useful genetic markers for analysis of genetic diversity, varietal identification, and phylogenetic relationships in peanut (*Arachis hypogaea*) cultivars and wild species. Genome 44:763–772
Rayburn AL, Biradar DP, Bullock DG et al (1997) Nuclear DNA content diversity in Chinese soybean introductions. Ann Bot 80:321–325
Resslar PM, Stucky JM, Miksche JP (1981) Cytophometric determination of the amount of DNA in *Arachis* (Leguminosae). Amer J Bot 68:149–153
Robledo G, Lavia GI, Seijo G (2009) Species relations among wild *Arachis* species with the A genome as revealed by FISH mapping of rDNA loci and heterochromatin detection. Theor Appl Genet 118:1295–1307
Robledo G, Lavia GI, Seijo G (2010) Genome re-assignment of *Arachis trinitensis* (Sect. Arachis, Leguminosae) and its implications for the genetic origin of cultivated peanut. Genet Mol Biol 33:714–718
Robledo G, Seijo G (2008) Characterization of the *Arachis* (Leguminosae) D genome using fluorescence in situ hybridization (FISH) chromosome markers and total genome DNA hybridization. Genet Mol Biol 31:717–724
Robledo G, Seijo G (2010) Species relationships among the wild B genome of *Arachis* species (section Arachis) based on FISH mapping of rDNA loci and

heterochromatin detection: a new proposal for genome arrangement. Theor Appl Genet 121:1033–1046
Samoluk SS, Chalup L, Robledo G et al (2015a) Genome sizes in diploid and allopolyploid *Arachis* L. species (section Arachis). Genet Resour Crop Evol 62:747–763
Samoluk SS, Robledo G, Podio M et al (2015b) First insight into divergence, representation and chromosome distribution of reverse transcriptase fragments from L1 retrotransposons in peanut and wild relative species. Genetica 143:113–125
Santana SH, Valls JFM (2015) *Arachis veigae* (Fabaceae), the most dispersed wild species of the genus, and yet taxonomically overloocked. Bonplandia 24:139–150
Schmidt T, Heslop- Harrison JS (1998) Genomes, genes and junk: the large-scale organization of plant chromosomes. Trends Plant Sci 3:195–199
Seijo G, Lavia GI, Fernández A et al (2004) Physical mapping of the 5S and 18S–25S rRNA genes by FISH as evidence that *Arachis duranensis* and *A. ipaensis* are the wild diploid progenitors of *A. hypogaea* (Leguminosae). Am J Bot 91:1294–1303
Seijo G, Lavia GI, Fernández A et al (2007) Genomic relationships between the cultivated peanut (*Arachis hypogaea*, Leguminosae) and its close relatives revealed by double GISH. Am J Bot 94:1963–1971
Shirasawa K, Koilkonda P, Aoki K et al (2012) *In silico* polymorphism analysis for the development of simple sequence repeat and transposon markers and construction of linkage map in cultivated peanut. BMC Plant Biol. doi:10.1186/147122291280
Silva SC, Martins MIG, Santos RC et al (2010) Karyological features and banding patterns in *Arachis* species belonging to the Heteranthae section. Plant Syst Evol 285:201–207
Silvestri MC, Ortiz AM, Lavia GI (2015) rDNA loci and heterochromatin positions support a distinct genome type for "x = 9 species" of section Arachis (*Arachis*, Leguminosae). Plant Syst Evol 301:555–562
Singh AK (1986) Utilization of wild relatives in the genetic improvement of *Arachis hypogaea* L. 8. Synthetic amphidiploids and their importance in interspecific breeding. Theor Appl Genet 72:433–439
Singh AK, Gurtu S, Jambunathan R (1993) Phylogenetic relationships in the genus *Arachis* based on seed protein profiles. Euphytica 74:219–225
Singh AK, Raina SN, Singh AK (1996) Variation in chromosomal DNA associated with the evolution of *Arachis* species. Genome 39:890–897
Singh AK, Simpson CE (1994) Biosystematics and genetic resources. In: Smartt J (ed) The groundnut crop, a scientific basis for improvement. Chapman and Hall, London, pp 96–138
Singh AK, Smartt J (1998) The genome donors of the groundnut/ peanut (*Arachis hypogaea* L.) revisited. Genet Resour Crop Evol 45:113–118
Slamovits CH, Cook JA, Lessa EP et al (2001) Recurrent amplifications and deletions of satellite DNA accompanied chromosomal diversification in South American Tuco-tucos (Genus *Ctenomys*, Rodentia: Octodontidae): a phylogenetic approach. Mol Biol Evol 18:1708–1719
Smartt J, Gregory WC (1967) Interspecific cross-compatibility between the cultivated peanut *Arachis hypogaea* L. and other members of the genus *Arachis*. Oléagineux 22:455–459
Smartt J, Gregory WC, Gregory MP (1978) The genome of *A. hypogaea*. 1. Cytogenetic studies of putative genome donors. Euphytica 27:665–675
Smartt J, Stalker HT (1982) Speciation and cytogenetics of *Arachis*. In: Pattee HE, Young CT (eds) Peanut science and technology. Peanut Res and Educ Soc, Yoakum, pp 21–49
Stalker HT (1991) A new species in section Arachis of peanuts with a D genome. Amer J Bot 78:630–637
Stalker HT, Dhesi JS, Parry DC et al (1991) Cytological and interfertility relationships of *Arachis* section Arachis. Am J Bot 78:238–246
Stalker HT, Moss JP (1987) Speciation, cytogenetics, and utilization of *Arachis* species. Adv Agron 41:1–40
Stalker HT, Wynne JC (1979) Cytology of interspecific hybrids in section Arachis of peanuts. Peanut Sci 6:110–114
Tallury SP, Hilu KW, Milla R et al (2005) Genomic affinities in *Arachis* section Arachis (Fabaceae): molecular and cytogenetic evidence. Theor Appl Genet 111:1229–1237
Temsch EM, Greilhuber J (2000) Genome size variation in *Arachis hypogaea* and *A. monticola* re-evaluated. Genome 43:449–451
Temsch EM, Greilhuber J (2001) Genome size in *Arachis* duranensis: a critical study. Genome 44:826–830
Urdampilleta JD, Pereira de Souza A, Schneider DRS et al (2009) Molecular and cytogenetic characterization of an AT-rich satellite DNA family in *Urvillea chacoensis* Hunz. (Paullinieae, Sapindaceae). Genetica 136:171–177
Valls JFM, Costa LC, Custodio AR (2013) A novel trifoliolate species of *Arachis* (Fabaceae) and further comments on the taxonomic section Trierectoides. Bonplandia 22:91–97
Valls JFM, Simpson CE (2005) New species of *Arachis* (Leguminosae) from Brazil, Paraguay and Bolivia. Bonplandia 14:35–63
Wendel JF, Cronn RC, Johnston JS et al (2002) Feast and famine in plant genomes. Genetica 115:37–47
Zhang L, Xu C, Yu W (2012) Cloning and Characterization of Chromosomal Markers from a Cot-1 Library of Peanut (*Arachis hypogaea* L.). Cytogenet Genome Res 137:31–41
Zhang L, Yang X, Tian L et al (2016) Identification of peanut (*Arachis hypogaea*) chromosomes using a fluorescence in situ hybridization system reveals multiple hybridization events during tetraploid peanut formation. New Phytol 211:1424–1439